

**PERBANDINGAN 3 METODE DALAM *DATA MINING* UNTUK  
PREDIKSI PENERIMA BEASISWA BERDASARKAN PRESTASI  
DI SMA NEGERI 6 SURAKARTA**



**SKRIPSI**

Disusun sebagai salah satu syarat menyelesaikan Jenjang Strata I

Pada Program Studi Informatika Fakultas Komunikasi dan Informatika

Universitas Muhammadiyah Surakarta

**Oleh :**

**VERONICA ANDRIYANA**

**L200110086**

**PROGRAM STUDI INFORMATIKA  
FEKULTAS KOMUNIKASI DAN INFORMATIKA  
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

**2015**

## **HALAMAN PERSETUJUAN**

Skripsi dengan judul

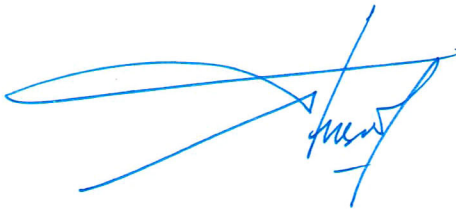
**PERBANDINGAN 3 METODE DALAM *DATA MINING* UNTUK  
PREDIKSI PENERIMA BEASISWA BERDASARKAN PRESTASI  
DI SMA NEGERI 6 SURAKARTA**

Ini telah diperiksa dan disetujui dalam sidang pendadaran :

Hari : JUMAT

Tanggal : 6 MARET 2015

Pembimbing



**Yusuf Sulisty Nugroho, S.T., M.Eng**

**NIK : 1197**

**HALAMAN PENGESAHAN**

**PERBANDINGAN 3 METODE DALAM *DATA MINING* UNTUK  
PREDIKSI PENERIMA BEASISWA BERDASARKAN PRESTASI  
DI SMA NEGERI 6 SURAKARTA**

Yang dipersiapkan dan disusun oleh :

**VERONICA ANDRIYANA**

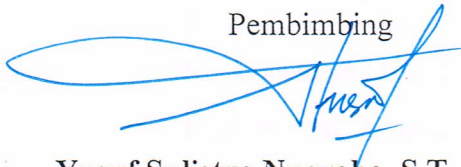
NIM : L200110086

Telah dipertahankan didepan Dewan Penguji

Pada tanggal 6 Maret 2015

**Susunan Dewan Penguji**

Pembimbing



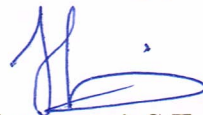
Yusuf Sulistyo Nugroho, S.T., M.Eng  
NIK : 1197

Dewan Penguji I



Fatah Yasin Irsyadi, S.T., M.T  
NIK : 738

Dewan Penguji II



Hasyim Asyari, S.T., M.T  
NIK : 981

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar sarjana

Tanggal 6 Maret 2015

Dekan

Fakultas Komunikasi dan Informatika



Husni Thamrin, S.T., M.T., Ph.D  
NIK 706

Ketua Program Studi  
Teknik Informatika



Dr. Heru Supriyono, M.Sc.  
NIK 970

## DAFTAR KONTRIBUSI

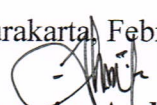
Dengan ini saya menyatakan bahwa skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Berikut saya sampaikan daftar kontribusi dalam penyusunan skripsi :

1. Yusuf Sulisty Nugroho, S.T., M.Eng, selaku dosen pembimbing yang memberikan masukan dan arahan dalam penyusunan skripsi ini.
2. Saya sendiri yang mengolah data dan perhitungan perbandingan 3 metode (*Naive Bayes*, *Decision Tree* Algoritma ID3, dan Regresi Linear) berdasarkan sumber – sumber pada buku, literatur dan internet.
3. Aplikasi yang saya gunakan untuk membantu skripsi ini yaitu :
  - a. *RapidMiner 5*
  - b. Microsoft Excell 2010
4. Laptop yang digunakan memiliki spesifikasi Processor Intel® Core™ i3 – 380M, RAM 2,00 GB, 320 HDD, Acer Nplify™ 802.11b/g/n

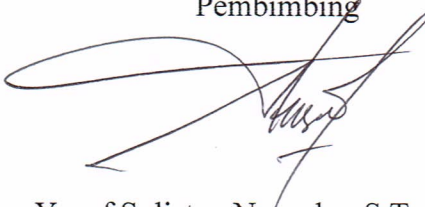
Dengan pernyataan dan daftar kontribusi ini saya buat dengan sejujurnya. Saya bertanggungjawab atas isi dan kebenaran daftar diatas.

Surakarta, Februari 2015

  
**Veronica Andriyana**

Mengetahui :

Pembimbing



Yusuf Sulisty Nugroho, S.T., M.Eng

NIK : 1197

## MOTTO DAN PERSEMBAHAN

### MOTTO :

*“ Man Jadda Wajada, siapa yang bersungguh – sungguh pasti akan berhasil*

*Man Shobaru Zhafira, siapa yang bersabar akan beruntung*

*Man Yasro Yahsud, siapa yang menanam, akan menuai seperti yang ditanam”*

- Ibunda -

*“Hidup adalah tentang proses, bukan tentang target akhir apa yang akan seseorang itu dapatkan. Tetapi tentang bagaimana, apa, dimana, siapa, dan mengapa untuk dapat menuju target tujuan tersebut”*

- My Beloved -

*“Jangan pernah takut untuk mencoba tentang hal yang kamu anggap sulit atau tidak bisa. Selalu coba dan coba, maka kamu akan mendapatkan kesuksesanmu”*

- Penulis -

## **PERSEMBAHAN :**

Dengan mengucapkan syukur Alhamdulillah atas rahmad dan hidayah Allah SWT, akan saya persembahkan karya ini kepada orang-orang yang saya sayangi :

1. Ayahanda Joko Setiawan dan Ibunda Sri Suranti yang tak pernah lelah dalam memotivasi, mendukung, dan mendoakan untuk meraih kesuksesan dunia dan akhirat.
2. Danni Andriawan Nugroho sebagai adik yang menghibur, terkadang mengantar saya, dan juga teman saat dirumah.
3. Setiawan Dwi Cahyo, seseorang yang selalu tulus menemaniku dalam memberikan semangat, memotivasi, dan menghiburku disaat aku merasa malas dan jenuh dalam menyelesaikan skripsi.
4. Syarifah Nur Haryati, sahabat terbaik yang selalu ada dan kompak dalam berjuang semasa kuliah, penasihat yang baik serta partner yang baik yang bersama – sama menyelesaikan skripsi dan akhirnya bisa lulus bersama.
5. Pasukan Damin (Syarifah dan Deny) sebagai partner yang berjuang bersama dan bertukar pikiran, panas dan hujan untuk menyelesaikan skripsi. Kalian luar biasa !!!
6. Irma, Novi, mas Ibnu, mbak Tika, yang membantu dan memberi dukungan.
7. Aldiana Putri, Riska Ayu, Yohanes, Akbar, Rizky, sahabat yang selalu menghibur disaat saya sedang suntuk dan menyemangati saya.
8. Rekan HIMATIF UMS, tempat pertama kali saya mengenal organisasi.
9. Teman – teman seperjuangan Teknik Informatika angkatan 2011 semuanya terima kasih telah berbagi ilmu dan pengalamannya.

## KATA PENGANTAR

Assalamu'alaikum warohmatullohi Wabarakatuh

Puji syukur kehadiran Allah SWT, yang telah melimpahkan nikmat dan anugerah yang sangat luar biasa sehingga saya dapat menyelesaikan skripsi dengan baik yang berjudul “Perbandingan 3 Metode Dalam *Data Mining* Untuk Prediksi Penerima Beasiswa Berdasarkan Prestasi Di SMA Negeri 6 Surakarta”.

Skripsi ini disusun sebagai syarat wajib untuk menyelesaikan pendidikan jenjang Strata 1 pada Program Studi Teknik Informatika Universitas Muhammadiyah Surakarta. Atas selesainya skripsi ini tidak lepas dari dukungan dan bantuan dari berbagai pihak sehingga dapat berjalan dengan lancar dan sesuai harapan, oleh karena itu pada kesempatan ini penulis ingin menyampaikan banyak terima kasih kepada :

1. Allah SWT yang telah memberikan nikmat dan hidayahnya kepada penulis sehingga diizinkan untuk menyelesaikan skripsi ini.
2. Orang tua dan keluarga besar penulis yang telah memberikan segala dukungan berupa doa, semangat, dan motivasi dengan tiada hentinya kepada penulis.
3. Bapak Husni Thamrin, S.T, M.T., Ph.D selaku dekan Fakultas Komunikasi Dan Informatika Universitas Muhammadiyah Surakarta.
4. Bapak Dr. Heru Supriyono, M.Sc selaku ketua Program Studi Teknik Informatika.

5. Bapak Yusuf Sulisty Nugroho, S.T., M.Eng selaku pembimbing akademik selama kuliah serta pembimbing skripsi yang telah memberikan bimbingan dan pengarahan kepada penulis sehingga dapat menyelesaikan tugas akhir ini.
6. Segenap dosen dan karyawan prodi Teknik Informatika atas bantuan dan ilmu yang diberikan kepada penulis selama masa perkuliahan hingga dinyatakan mendapat gelar Strata 1.
7. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang berperan dan membantu dalam menyelesaikan skripsi ini.

Semoga Allah SWT selalu melimpahkan nikmat dan hidayahNya kepada orang – orang yang telah membantu dalam kelancaran penyusunan ini. Penulis berharap semoga skripsi ini dapat bermanfaat bagi penulis dan orang lain untuk dikembangkan sehingga akan dapat menghasilkan penelitian – penelitian yang lebih baik lagi dari sebelum – sebelumnya.

Wassalamu’alaikum Warohmatullohi Wabarokatuh.

Surakarta, Februari 2015

Penulis



## DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
DAFTAR KONTRIBUSI.....	iv
MOTTO DAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR PERSAMAAN .....	xv
DAFTAR LAMPIRAN .....	xvi
ABSTRAKSI .....	xvii
<b>BAB I     PENDAHULUAN.....</b>	<b>1</b>
A. Latar Belakang Masalah .....	1
B. Perumusan Masalah .....	3
C. Batasan Masalah .....	3
D. Tujuan Penelitian .....	5
E. Manfaat Penelitian .....	5
F. Sistematika Penulisan .....	7
<b>BAB II    TINJAUAN PUSTAKA .....</b>	<b>8</b>
A. Telaah Penelitian .....	8

B. Landasan Teori .....	13
1. Prediksi / Peramalan.....	13
2. <i>Data Mining</i> .....	14
3. <b><i>Naive Bayes</i></b> .....	<b>15</b>
4. <i>Decision Tree</i> Algoritma ID3 .....	15
5. Regresi Linear .....	16
6. <i>RapidMiner 5</i> .....	16
7. <i>Precision</i> .....	17
8. <i>Recall</i> .....	17
9. <i>Accuracy</i> .....	17
<b>BAB III   METODOLOGI PENELITIAN</b> .....	<b>18</b>
A. Waktu dan Tempat Penelitian .....	18
B. Peralatan Utama dan Pendukung.....	18
C. Langkah Penelitian.....	19
1. Identifikasi Masalah .....	19
2. Metode Pengumpulan Data .....	19
3. Menentukan Kebutuhan .....	21
4. Pengelompokan Data.....	24
5. Implementasi Data Mining.....	26
1. <i>Naive Bayes</i> .....	27
2. <i>Decision Tree</i> Algoritma ID3 .....	27
3. Regresi Linear .....	28

<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>30</b>
A.	Hasil Penelitian .....	30
1.	Menentukan Atribut .....	30
2.	Pengolahan Data.....	31
3.	Hasil Implementasi <i>Naive Bayes</i> menggunakan <i>RapidMiner 5</i> .....	32
4.	Hasil Implementasi <i>Decision Tree</i> Algoritma ID3 .....	35
5.	Hasil Implementasi Regresi Linear menggunakan <i>RapidMiner 5</i> .....	43
B.	Analisa dan Pembahasan .....	45
1.	Penentuan Sampel .....	45
2.	Implementasi dengan Perhitungan <i>Naive Bayes</i> .....	46
3.	Implementasi dengan Perhitungan <i>Decision Tree</i> .....	51
4.	Implementasi dengan Perhitungan Regresi Linear .....	96
C.	Perbandingan 3 Metode .....	100
D.	Analisis Penelitian .....	101
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN .....</b>	<b>108</b>
A.	Kesimpulan .....	108
B.	Saran .....	110
<b>DAFTAR PUSTAKA .....</b>		<b>111</b>
<b>LAMPIRAN .....</b>		<b>114</b>

## DAFTAR TABEL

Tabel 3.1	Data Variabel awal.....	22
Tabel 3.2	Data Variabel yang digunakan .....	23
Tabel 3.3	Pengelompokkan data .....	25
Tabel 4.1	Data <i>Training</i> .....	31
Tabel 4.2	Data Variabel Regresi Linear.....	102
Tabel 4.3	Perbandingan metode <i>Naive Bayes</i> , <i>Decision Tree</i> Algoritma ID3, dan Regresi Linear.....	105
Tabel 4.4	Kuesioner untuk Siswa.....	106
Tabel 4.3	Kuesioner untuk Guru .....	109

## DAFTAR GAMBAR

Gambar 4.1	Rancangan proses <i>Naive Bayes</i> menggunakan data <i>training</i> .	33
Gambar 4.2	Rancangan proses <i>Naive Bayes</i> menggunakan data <i>testing</i> ...	34
Gambar 4.3	Hasil prediksi penerima <i>beasiswa</i> .....	35
Gambar 4.4	Tampila <i>hasil Naive Bayes</i> pada <i>Text View</i> .....	36
Gambar 4.5	<i>Scatter Plot Naive Bayes</i> menggunakan data <i>training</i> .....	36
Gambar 4.6	<i>Scatter Plot Naive Bayes</i> menggunakan data <i>testing</i> .....	38
Gambar 4.7	Rancangan proses penerapan <i>Decision Tree</i> untuk prediksi penerima <i>beasiswa</i> .....	39
Gambar 4.8	Hasil skema <i>Decision Tree</i> untuk penerima <i>beasiswa</i> .....	40
Gambar 4.9	Tampilan hasil <i>Decision Tree</i> pada <i>Scatter Plot</i> .....	46
Gambar 4.10	Rancangan Proses Regresi Linear.....	47
Gambar 4.11	Hasil Regresi Linear.....	47
Gambar 4.12	<i>Scatter view</i> Regresi Linear .....	48
Gambar 4.13	<i>Root Node</i> .....	64
Gambar 4.14	<i>Internal Node</i> ke – 2 .....	71
Gambar 4.15	<i>Internal Node</i> ke – 2 .....	76
Gambar 4.16	<i>Leaf node</i> ekstrakurikuler Olahraga.....	80
Gambar 4.17	<i>Internal node Gender</i> .....	85
Gambar 4.18	<i>Leaf node gender</i> Laki – laki .....	89
Gambar 4.19	<i>Internal node</i> Gaji orang tua .....	93
Gambar 4.20	<i>Internal node</i> Gaji orang tua sebesar gaji $\leq 1500000$ .....	95

Gambar 4.21	<i>Leaf node</i> Jumlah tanggungan orang tua.....	97
Gambar 4.22	Pohon Keputusan berdasarkan Nilai Rata – rata $6 < \text{nilai} \leq 8$ ...	100
Gambar 4.23	Grafik Kuesioner Siswa .....	106
Gambar 4.24	Grafik Kuesioner Guru .....	109

## DAFTAR PERSAMAAN

Persamaan 3.1 Rumus <i>Slovin</i> .....	24
Persamaan 3.2 Rumus Teorema <i>Bayes</i> .....	26
Persamaan 3.3 Rumus <i>Entropy</i> .....	28
Persamaan 3.4 Rumus <i>Information Gain</i> .....	28
Persamaan 3.5 Rumus Regresi Linear Sederhana untuk Populasi .....	29
Persamaan 3.6 Rumus Pendekatan Regresi Linear Sederhana .....	29
Persamaan 3.7 Rumus nilai optimal $b_1$ .....	29
Persamaan 3.8 Rumus nilai optimal $b_0$ .....	29
Persamaan 3.9 Rumus Perhitungan manual Regresi Linear.....	29

## DAFTAR LAMPIRAN

Lampiran 1. Data *Training*

Lampiran 2. Data *Testing*

Lampiran 3. Hasil Prediksi dengan menggunakan metode *Naive Bayes*

Lampiran 4. Tampilan *Graph View Decision Tree* Algoritma ID3 hasil Prediksi

Penerima Beasiswa dengan menggunakan *RapidMiner 5*

Lampiran 5. Hasil Kuesioner

Lampiran 6. Aturan terkait beasiswa



## ABSTRAKSI

Dalam rangka meningkatkan akses dan minat belajar siswa serta mengangkat mutu sekolah, SMA N 6 Surakarta mengalokasikan dana beasiswa dalam bentuk apresiasi untuk siswa berprestasi. Namun masih ada hal yang menjadi permasalahan yang sering muncul, yaitu kurang tepatnya penyaluran beasiswa terhadap siswa. Beasiswa untuk siswa berprestasi bertujuan memotivasi siswa untuk selalu meningkatkan prestasi akademik maupun non akademik dan membantu siswa yang kurang mampu tetapi berprestasi. Untuk mengatasi permasalahan tersebut adalah dengan cara menerapkan proses *data mining*.

Dalam memprediksi siswa yang menerima beasiswa berdasarkan prestasi menggunakan metode *Naive Bayes*, *Decision Tree* Algoritma ID3, dan Regresi Linear. Atribut yang digunakan terdiri dari Nilai rata-rata, *Gender*, Ekstrakurikuler, Jurusan, Semester, Jumlah Tanggungan Orang Tua, Gaji Orang Tua, dan Beasiswa. Untuk melakukan proses *data mining* tersebut di perlukan tools pembantu yaitu *RapidMiner 5*.

Pengimplementasian *data mining* menggunakan perbandingan 3 metode dapat diketahui bahwa berdasarkan dari jumlah sampel 305 siswa hasil nilai *precision* metode *Decision Tree* Algoritma ID3 lebih baik digunakan untuk penelitian ini dibandingkan dengan metode yang lain. Sedangkan berdasarkan nilai *recall* dan *accuracy*, Regresi Linear lebih baik digunakan dibandingkan metode lain. Tetapi apabila dilihat dari hasil secara keseluruhan prediksi penerima beasiswa variabel yang paling berpengaruh adalah Nilai rata-rata.

**Kata kunci :** *Algoritma ID3, Data mining, Decision Tree, Naive Bayes, Regresi Linear*